

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 2 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

LISTING OF CLAIMS

This listing of claims replaces all prior versions and listings of claims in the application:

1. (Currently Amended) A method, comprising:

~~defining receiving a definition of at least one a~~ target rule for detection of target hits in an article, ~~wherein the target rule comprising comprises:~~

~~defining a concept list comprising an origin concept, a distance, and a relationship between the origin concept and an evaluated term, and a distance representing a strength of the relationship between the origin concept and the evaluated term, and~~

~~defining a target scope that characterizes a document region to which the target rule is to be applied;[;]~~

~~defining receiving a definition of at least one an~~ extraction rule ~~that is based on the at least one target rule, wherein the extraction rule definition comprising comprises~~ defining an extraction scope ~~that characterizes a document region to be extracted;~~

~~applying the at least one target rule to each document regions target-scope of the article characterized by the target scope;~~

~~determining a target score for each target-scope the document regions of the article, wherein the score represents how well the target rule is satisfied by the document regions;~~

~~applying the at least one extraction rule to the article to determine at least one an extract from the article, wherein the application of the extraction rule is based on at least one the determined target score; and~~

~~outputting the at least one extract.~~

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 3 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

2. (Currently Amended) The method of claim 1, wherein:
applying the target rule comprises detecting a plurality of target hits are detected; and
applying the extraction rule comprises extracting a plurality of extracts are extracted.
3. (Currently Amended) The method of claim 2, further comprising sorting the extracts based on the ~~at least one~~ extraction rule.
4. (Currently Amended) The method of claim 1, further comprising ~~extracting at least~~
selecting one a first extract from the article for output based on the target score.
5. (Currently Amended) The method of claim 1, wherein the ~~at least one~~ target rule further comprises a target score formula for determining the target score.
6. (Currently Amended) The method of claim 5, wherein applying the ~~at least one~~ target rule comprises using the target score formula to detect target hits.
7. (Currently Amended) The method of claim 1, wherein the ~~concept~~ target rule further comprises a ~~concept set~~, a gist or both defined as a vector of weighted concepts.
8. (Currently Amended) The method of claim ~~[[7]]~~ 1, wherein the target rule further comprises a concept set that comprises a list of concepts the concept list and a second concept.
9. (Currently Amended) The method of claim 8, wherein ~~the second concepts in the list of concepts~~ concept are produced by comprises a product of set operations on multiple lists of
two or more other concepts.

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 4 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

10. (Currently Amended) The method of claim 7, wherein ~~the~~ the gist comprises weighted concepts a user defined gist.

11. (Currently Amended) The method of claim 7, wherein ~~the gist is user-defined or is~~ comprises a calculated gist of the article a document region .

12. (Currently Amended) The method of claim 1, wherein ~~the document region~~ characterized by the target scope comprises an article, a sentence, or a term.

13. (Currently Amended) The method of claim 1, wherein ~~the document region~~ characterized by the extraction scope comprises an article, a sentence, or a term.

14. (Currently Amended) The method of claim 1, further comprising preprocessing the article, wherein preprocessing comprises:

determining at least one concept associated with the article~~;~~[]; and

determining a gist of the article.

15. (Currently Amended) ~~A~~ An article comprising one or more computer-readable medium media containing program code operable to cause one or more machines to perform operations. the operations comprising:

~~program-code-for-defining~~ receiving a definition of at least one a target rule for detection of target hits in an article, wherein the target rule comprising comprises :

defining a concept list comprising an origin concept, a distance, and a relationship between the origin concept and an evaluated term, and a distance representing a strength of the relationship between the origin concept and the evaluated term, and

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 5 of 12

Attorney's Docket No : 16113-330001 / GP-138-00-US

defining a target scope that characterizes a document region to which the target rule is to be applied; [[:]]

program-code-for defining receiving a definition of at-least-one an extraction rule that is based on the at-least-one target rule, wherein the extraction rule definition comprising comprises defining an extraction scope that characterizes a document region to be extracted;

program-code-for applying the at-least-one target rule to each document regions target scope of the article characterized by the target scope;

program-code-for determining a target score for each target scope the document regions of the article, wherein the score represents how well the target rule is satisfied by the document regions;

program-code-for applying the at-least-one extraction rule to the article to determine at least-one an extract from the article, wherein the application of the extraction rule is based on at least-one the determined target score; and

program-code-for outputting the at-least-one extract.

16. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, wherein: applying the target rule comprises detecting a plurality of target hits are detected ; and applying the extraction rule comprises extracting a plurality of extracts are extracted..

17. (Currently Amended) The ~~computer-readable medium~~ article of claim 16, further comprising program-code-for sorting the extracts based on the at-least-one extraction rule.

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 6 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

18. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, further comprising ~~program code for extracting at least~~ selecting one a first extract from the article for output based on the target score.

19. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, wherein the ~~at least one~~ target rule further comprises a target score formula for determining the target score.

20. (Currently Amended) The ~~computer-readable medium~~ article of claim 19, wherein applying the ~~at least one~~ target rule comprises using the target score formula to detect target hits.

21. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, wherein the ~~concept~~ target rule further comprises a concept set; a gist or ~~both~~ defined as a vector of weighted concepts.

22. (Currently Amended) The ~~computer-readable medium~~ article of claim ~~[[21]]~~ 15, wherein the target rule further comprises a concept set that comprises a list of concepts the concept list and a second concept.

23. (Currently Amended) The ~~computer-readable medium~~ article of claim 22, wherein the second concepts in the list of concepts concept are produced by comprises a product of set operations on multiple lists of two or more other concepts.

24. (Currently Amended) The ~~computer-readable medium~~ article of claim 21, wherein a the gist comprises weighted concepts a user defined gist.

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 7 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

25. (Currently Amended I) The ~~computer-readable medium~~ article of claim 21 wherein the gist is ~~user-defined or is~~ comprises a calculated gist of the ~~article~~ a document region.

26. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, wherein the document region characterized by the target scope comprises an article, a sentence, or a term.

27. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, wherein the document region characterized by the extraction scope comprises an article, a sentence, or a term.

28. (Currently Amended) The ~~computer-readable medium~~ article of claim 15, further comprising ~~program code for~~ preprocessing the article, wherein preprocessing comprises:
determining at least one concept associated with the article[]; and
determining a gist of the article.

29. (Currently Amended) The method of claim 1, wherein:
the origin concept ~~further~~ comprises a group of related words, relationships with other concepts, the ~~strength~~ strengths of the relationships, and statistics about regarding the ~~concept~~ usage of the origin concept in a language.

30. (Currently Amended) A method, comprising:
defining receiving a definition of a target rules rule for detection of one or more target hits in an article, wherein the target rule definition comprising comprises defining a definition of a target article region, a comparison method, and a target definition comprising a relationship, a distance, and at least one of a concept, a concept set, or a gist, a relationship between the at least

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 8 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

one concept, concept set, or gist and an evaluated term, and a distance representing a strength of the relationship;

defining receiving a definition of an extraction rules rule for the extraction of one or more extracts from the article, wherein the extraction rule comprising comprises defining a definition of an extraction article region and a description of the at least one target rule definition, wherein the target rule relates to which the extraction rules rule relate;

applying the target rules rule to each-target a collection of article region regions of the article to determine a target score for each-target the article region regions; and

applying the extraction rules rule to detect identify at least one extract associated with at least one target article region from the article based at least in part on the target score scores for the at least one target article region regions; and

outputting the at least one extract.

31. (Currently Amended) The method of claim 30, further-comprising wherein:

applying the extraction rule comprises determining a plurality of extracts based on the extraction rules; and

the method further comprises ranking the extracts.

32. (Currently Amended) The method of claim 30, wherein the gist is comprises a vector of weighted concepts.

33. (Previously Presented) The method of claim 30, wherein the concept is defined by a group of related words, relationships with other concepts, the strength of the relationships, and statistics about the concept usage in language.

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 9 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

34. (Currently Amended) The method of claim 30, wherein the comparison method is comprises one of an associated categorization-like method, an inverse categorization-like method, a similarity to gist method, or a concept set hit method.

35. (Currently Amended) The method of claim 1, wherein defining receiving the definition of the at least one extraction rule further comprises defining receiving a definition of a sort order in which extracts are to be sorted for output.

36. (Currently Amended) The ~~method~~ article of claim 15, wherein defining receiving the definition of the at least one extraction rule further comprises defining receiving a definition of a sort order in which extracts are to be sorted for output.

37. (Currently Amended) The method of claim 1, wherein the distance comprises a numeric represents representation of the [[a]] strength of the relationship between the origin concept and an evaluated concept.

38. (Previously Presented) The method of claim 1, wherein the relationship comprises one of "is a product of", "is a part of", "is a kind of", "has kind", or "has part".

39. (Previously Presented) The method of claim 1, wherein the origin concept comprises at least one search term.

40. (Previously Presented) The method of claim 8, wherein the concept set further comprises at least one set operation.

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 10 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

41. (Previously Presented) The method of claim 40, wherein the set operation comprises one of "AND", "OR", and "AND NOT".

42. (Currently Amended) The ~~method~~ article of claim 15, wherein the distance comprises a numeric represents representation of the ~~[[a]]~~ strength of the relationship between the origin concept and an evaluated concept.

43. (Currently Amended) The ~~method~~ article of claim 15, wherein the relationship comprises one of "is a product of", "is a part of", "is a kind of", "has kind", or "has part".

44. (Currently Amended) The ~~method~~ article of claim 15, wherein the origin concept comprises at least one search term.

45. (Currently Amended) The ~~method~~ article of claim 22, wherein the concept set further comprises at least one set operation.

46. (Currently Amended) The ~~method~~ article of claim 45, wherein the set operation comprises one of "AND", "OR", and "AND NOT".

47. (New) A method for extracting a subset of a document, comprising:
receiving information describing two or more concepts, wherein a concept is defined through a collection of related words;
combining the two or more concepts using an operation to define a target definition that is to be detected;
receiving a description of a document region targeted for extraction;
accessing a document;

Applicant : Weissman et al.
Serial No. : 10/741,303
Filed : December 18, 2003
Page : 11 of 12

Attorney's Docket No.: 16113-330001 / GP-138-00-US

based on the target definition and the document regions targeted for extraction, extracting one or more regions of the accessed document; and
making the extracted regions available for output.

48. (New) The method of claim 47, wherein the concepts each comprise a lexical concept defined by a group of related words and relationships with related concepts.